Best Linear Unbiased Prediction (BLUP)

C. R. Henderson

- Born April 1, 1911, in Coin, Iowa, in Page County – same county of birth as Jay Lush
- Page County Farm Bureau Picnic (12 and under, 14 and under, 16 and under changed to 10-12, 13-14, 15-16)
- Dean H.H. Kildee visited in 1929 and convinced Henderson to come to Iowa State College.

C. R. Henderson

- Returned to ISU after the war for Ph.D. with Jay Lush in animal breeding.
- Professor at Cornell until 1976
- Renowned for “Henderson’s Mixed Model Equations” and work on BLUP.
- Elected member of the National Academy of Sciences

C. R. Henderson

- ISC track 4 x 220 indoor world record
- 1933 ISC Field House indoor 440 record of 51.7 (stood for 30 years)
- Outdoor 440 record of 48.6 when world record was 47.4
- MS in nutrition from ISC
Henderson’s Ph.D. Students Included

• Shayle Searle (who taught Henderson matrix algebra)

• David Harville (professor emeritus, Department of Statistics, ISU, linear models expert)

Henderson’s Advice to Young Scientists

1. Study methods of your predecessors.
2. Work hard.
3. Do not fear to try new ideas.
4. Discuss your ideas with others freely.
5. Be quick to admit errors. Progress comes by correcting mistakes.
6. Always be optimistic. Nature is benign.
7. Enjoy your scientific work. It can be a great joy.

Sources


A problem that reportedly sparked Henderson’s interest in BLUP

We present here a variation of the original problem 23 on page 164 of

• Suppose intelligence quotients (IQs) for a population of students are normally distributed with a mean $\mu$ and variance $\sigma_u^2$. An IQ test was given to an i.i.d. sample of such students.

• Given the IQ of a student, the test score for that student is normally distributed with a mean equal to the student's IQ and a variance $\sigma_e^2$ and is independent of the test score of any other students.

• Suppose it is known that $\sigma_u^2 / \sigma_e^2 = 9$.

• If the sample mean of the students’ test scores was 100, what is the best prediction of the IQ of a student who scored 130 on the test?

Because $\mathbf{U}$ is a random vector rather than a fixed parameter, we talk about predicting $\mathbf{U}$ rather than estimating $\mathbf{U}$.

We seek a Best Linear Unbiased Predictor (BLUP) for $\mathbf{U}$, which we will denote by $\hat{\mathbf{U}}$.

Consider our linear mixed effects model

$$\mathbf{Y} = \mathbf{X} \beta + Z \mathbf{U} + \mathbf{e},$$

where

$$\begin{bmatrix} \mathbf{U} \\ \mathbf{e} \end{bmatrix} \sim N \left( \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \mathbf{0} \\ \mathbf{0} & \sigma_e^2 \end{bmatrix} \right).$$

Given data $\mathbf{Y}$, what is our best guess for the unobserved vector $\mathbf{U}$?

To be a BLUP, we require

1. $\hat{\mathbf{U}}$ is a linear function of $\mathbf{Y}$.

2. $\hat{\mathbf{U}}$ is unbiased for $\mathbf{U}$ so that $E(\hat{\mathbf{U}} - \mathbf{U}) = \mathbf{0}$.

3. $\text{Var}(\hat{\mathbf{U}} - \mathbf{U})$ is no “larger” than the $\text{Var}(\hat{\mathbf{U}} - \mathbf{U})$, where $\mathbf{X}$ is any other linear and unbiased predictor.
It turns out that the BLUP for $\mathbf{u}$ is the BLUE of $E(\mathbf{u}|\mathbf{y})$.

What does $E(\mathbf{u}|\mathbf{y})$ look like?

We will use the following result about conditional distributions for multivariate normal vectors.

Now note that

$$
\begin{bmatrix}
\mathbf{y} \\
\mathbf{u}
\end{bmatrix} = 
\begin{bmatrix}
\mathbf{X}\mathbf{b} \\
\mathbf{0}
\end{bmatrix} +
\begin{bmatrix}
\mathbf{Z} & \mathbf{I} \\
\mathbf{I} & \mathbf{0}
\end{bmatrix}
\begin{bmatrix}
\mathbf{u} \\
\mathbf{e}
\end{bmatrix}.
$$

Thus,

$$
\begin{bmatrix}
\mathbf{y} \\
\mathbf{u}
\end{bmatrix} \sim N
\left(
\begin{bmatrix}
\mathbf{X}\mathbf{b} \\
\mathbf{0}
\end{bmatrix},
\begin{bmatrix}
\mathbf{Z} & \mathbf{I} \\
\mathbf{I} & \mathbf{0}
\end{bmatrix}
\begin{bmatrix}
\mathbf{G} & \mathbf{0} \\
\mathbf{0} & \mathbf{I}
\end{bmatrix}
\begin{bmatrix}
\mathbf{Z}' \\
\mathbf{I}
\end{bmatrix}
\right).
$$

Suppose

$$
\begin{bmatrix}
\mathbf{w}_1 \\
\mathbf{w}_2
\end{bmatrix} \sim N
\left(
\begin{bmatrix}
\mathbf{u}_1 \\
\mathbf{u}_2
\end{bmatrix},
\begin{bmatrix}
\Sigma_1 & \Sigma_{12} \\
\Sigma_{21} & \Sigma_2
\end{bmatrix}
\right),
$$

where $\Sigma = \begin{bmatrix} \Sigma_1 & \Sigma_{12} \\ \Sigma_{21} & \Sigma_2 \end{bmatrix}$ is a positive definite variance matrix. Then the conditional distribution of $\mathbf{w}_2$ given $\mathbf{w}_1$ is as follows.

$$
(\mathbf{w}_2 | \mathbf{w}_1) \sim N
\left(
\mathbf{u}_2 + \Sigma_{21}\Sigma_{11}^{-1}(\mathbf{w}_1 - \mathbf{u}_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}
\right).
$$

Thus,

$$
E(\mathbf{u}|\mathbf{y}) = \mathbf{0} + \mathbf{G}\mathbf{Z}'(\mathbf{Z}\mathbf{G}\mathbf{Z}'+\mathbf{R})^{-1}(\mathbf{y}-\mathbf{X}\mathbf{b})
$$

$$
= \mathbf{G}\mathbf{Z}'\Sigma^{-1}(\mathbf{y}-\mathbf{X}\mathbf{b})
$$

Thus, the BLUP of $\mathbf{u}$ is

$$
\mathbf{G}\mathbf{Z}'\Sigma^{-1}(\mathbf{y}-\mathbf{X}\mathbf{b}_{\mathbf{z}})
$$

$$
= \mathbf{G}\mathbf{Z}'\Sigma^{-1}(\mathbf{y}-\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{y})
$$

$$
= \mathbf{G}\mathbf{Z}'\Sigma^{-1}(\mathbf{I}-\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1})\mathbf{y}.
$$
Suppose intelligence quotients (IQs) for a population of students are normally distributed with a mean $\mu$ and variance $\sigma^2$. An IQ test was given to an i.i.d. sample of such students.

Given the IQ of a student, the test score for that student is normally distributed with a mean equal to the student's IQ and a variance $\sigma^2$, and is independent of the test score of any other students.

If we let $\mu + \epsilon_i$, denote the IQ of student $i$ ($i = 1, \ldots, n$), then the test scores of the students are $N(\mu, \sigma^2)$ as in the statement of the problem.

If we let $Y_i = \mu + \epsilon_i + \varepsilon_i$, denote the test score of student $i$ ($i = 1, \ldots, n$), then the test scores of student $i$ are $N(\mu, \sigma^2 + \sigma^2)$ as in the problem statement.

If $\frac{\sum_{i=1}^{n} Y_i}{n}$ is the best prediction of the IQ of a student who scored 130 on the test, what is the sample mean of the students' test scores?
We have 
\[ \tilde{\mathbf{z}} = \mathbf{X} \mathbf{\beta} + \mathbf{Z} \mathbf{u} + \mathbf{\varepsilon}, \quad \text{where} \]
\[ \mathbf{X} = \mathbf{1}, \quad \mathbf{\beta} = \mathbf{u}, \quad \mathbf{Z} = \mathbf{I}, \quad \mathbf{\varepsilon} = \mathbf{\sigma}_u^2 \mathbf{I}, \quad \mathbf{R} = \mathbf{\sigma}_e^2 \mathbf{I}, \]
\[ \mathbf{\Sigma} = \mathbf{Z} \mathbf{Z}' + \mathbf{R} = (\mathbf{\sigma}_u^2 + \mathbf{\sigma}_e^2) \mathbf{I}. \]

Thus \[ \mathbf{\hat{\beta}} = (\mathbf{X}' \mathbf{\Sigma}^{-1} \mathbf{X})^{-1} \mathbf{X}' \mathbf{\Sigma}^{-1} \mathbf{y} \]
\[ = \begin{pmatrix} 1' \\ 1' \end{pmatrix} ^{-1} \mathbf{1}' \mathbf{y} = \mathbf{\bar{y}}. \quad \text{and} \]
\[ \mathbf{G} \mathbf{z} \mathbf{\Sigma}^{-1} = \frac{\mathbf{\sigma}_u^2}{\mathbf{\sigma}_u^2 + \mathbf{\sigma}_e^2} \mathbf{I}. \]

Thus, the BLUP for \( \mathbf{u} \) is
\[ \hat{\mathbf{u}} = \mathbf{G} \mathbf{z} \mathbf{\Sigma}^{-1} (\mathbf{y} - \mathbf{X} \hat{\mathbf{\beta}}) = \frac{\mathbf{\sigma}_u^2}{\mathbf{\sigma}_u^2 + \mathbf{\sigma}_e^2} (\mathbf{y} - \mathbf{1} \mathbf{\bar{y}}). \]

The \( i \)th element of this vector is
\[ \hat{\mathbf{u}}_i = \frac{\mathbf{\sigma}_u^2}{\mathbf{\sigma}_u^2 + \mathbf{\sigma}_e^2} (\mathbf{y}_i - \mathbf{\bar{y}}). \]

Thus, the BLUP for the IQ of student \( i \), \( \mathbf{u} + \hat{\mathbf{u}}_i \), is
\[ \hat{\mathbf{u}} + \hat{\mathbf{u}}_i = \mathbf{\bar{y}} + \frac{\mathbf{\sigma}_u^2}{\mathbf{\sigma}_u^2 + \mathbf{\sigma}_e^2} (\mathbf{y}_i - \mathbf{\bar{y}}) = \frac{\mathbf{\sigma}_u^2}{\mathbf{\sigma}_u^2 + \mathbf{\sigma}_e^2} \mathbf{y}_i + \frac{\mathbf{\sigma}_e^2}{\mathbf{\sigma}_u^2 + \mathbf{\sigma}_e^2} \mathbf{\bar{y}}. \]

Note that the BLUP is a convex combination of the individual score and the overall mean score.

\[ \frac{\mathbf{\sigma}_u^2}{\mathbf{\sigma}_u^2 + \mathbf{\sigma}_e^2} \mathbf{y}_i + \frac{\mathbf{\sigma}_e^2}{\mathbf{\sigma}_u^2 + \mathbf{\sigma}_e^2} \mathbf{\bar{y}}. \]

Because \( \mathbf{\sigma}_u^2 / \mathbf{\sigma}_e^2 \) is assumed to be 9, the weights are
\[ \frac{\mathbf{\sigma}_u^2}{\mathbf{\sigma}_u^2 + \mathbf{\sigma}_e^2} = \frac{9 \mathbf{\sigma}_u^2}{\mathbf{\sigma}_u^2 + 1} = \frac{9}{9 + 1} = 0.9 \]

and \( \frac{\mathbf{\sigma}_e^2}{\mathbf{\sigma}_u^2 + \mathbf{\sigma}_e^2} = 0.1 \). We would guess 0.9 (130) + 0.1 (100) = 127 to be the IQ of a student who scored 130 on the test.